

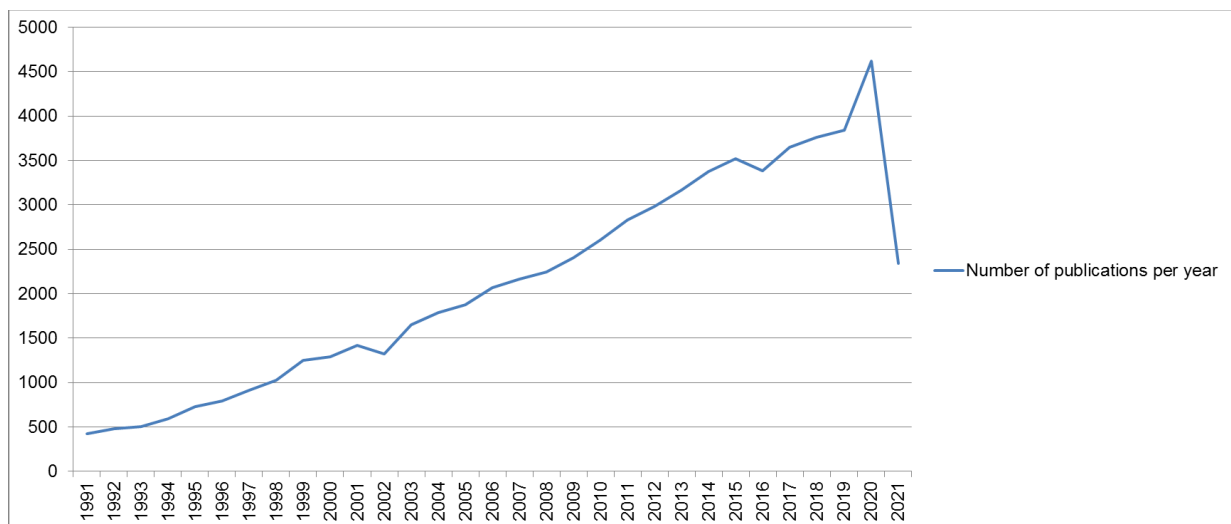
Supplementary Data S1

Stephen J. Goodswen, Paul J. Kennedy, John T. Ellis

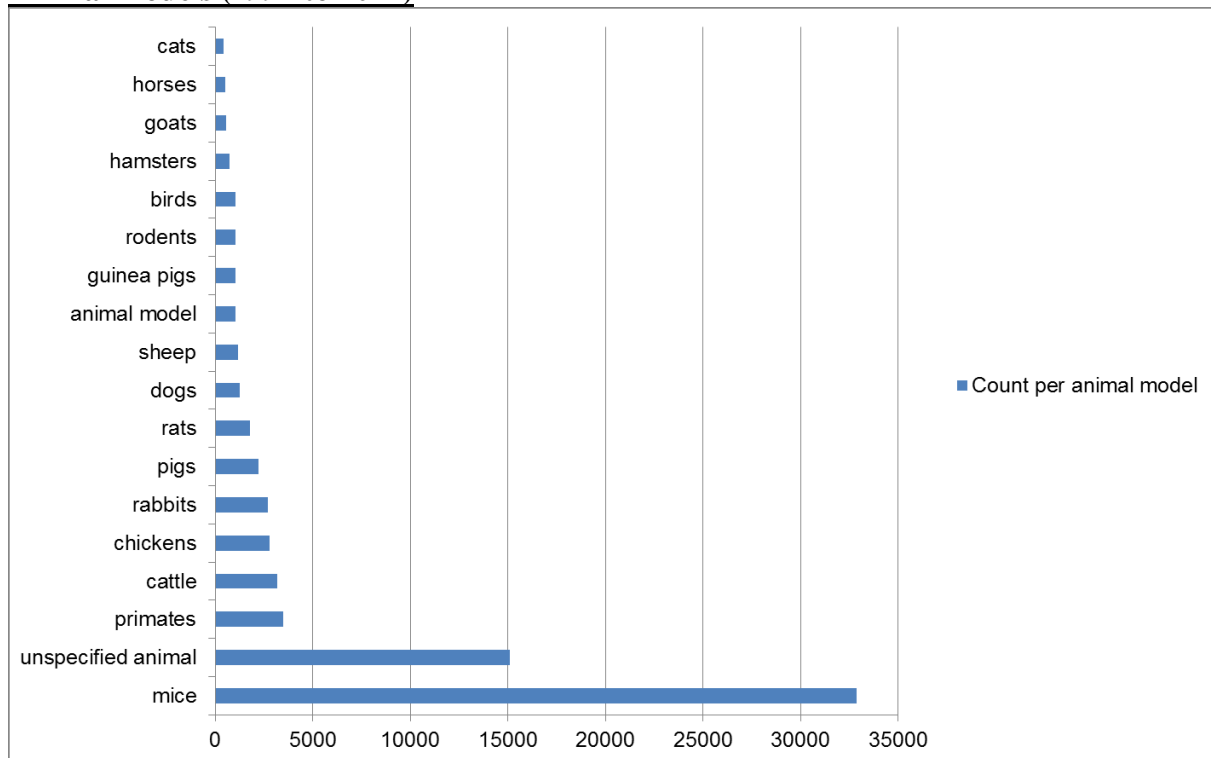
Introduction

This document contains graphs showing the counts (or frequency) of words relating to animal models, diseases caused by parasites, or the genus of the disease causing parasites that are found in PubMed derived 'title + abstract' text from publications. The 'title + abstract' text is the output from the machine learning abstract classification stage of the pipeline. Given an initial input of 332,627 'title + abstract' texts, 64,986 had a classification probability greater than or equal to 50% and were deemed 'abstracts of interest'. The words in these 'abstracts of interest' were used for the word counts.

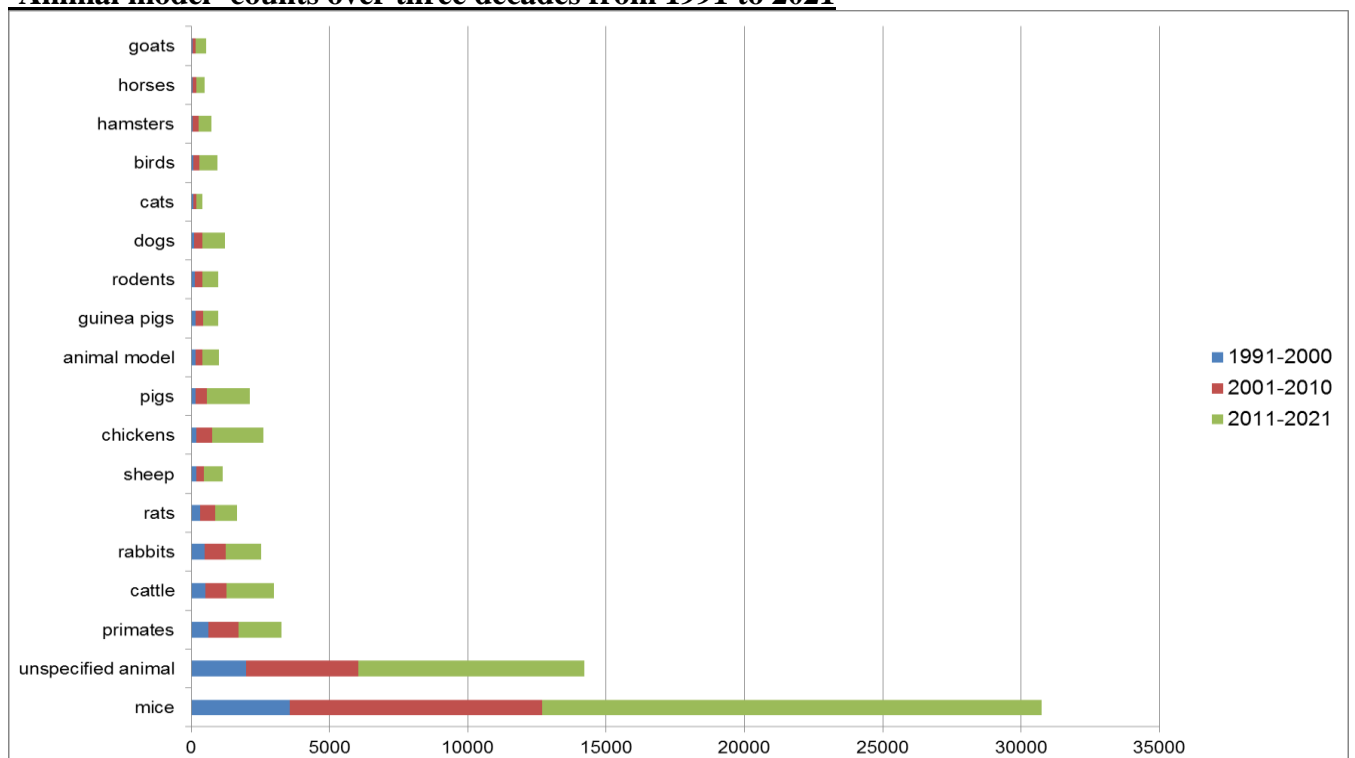
Number of publications from 1991 to 2021



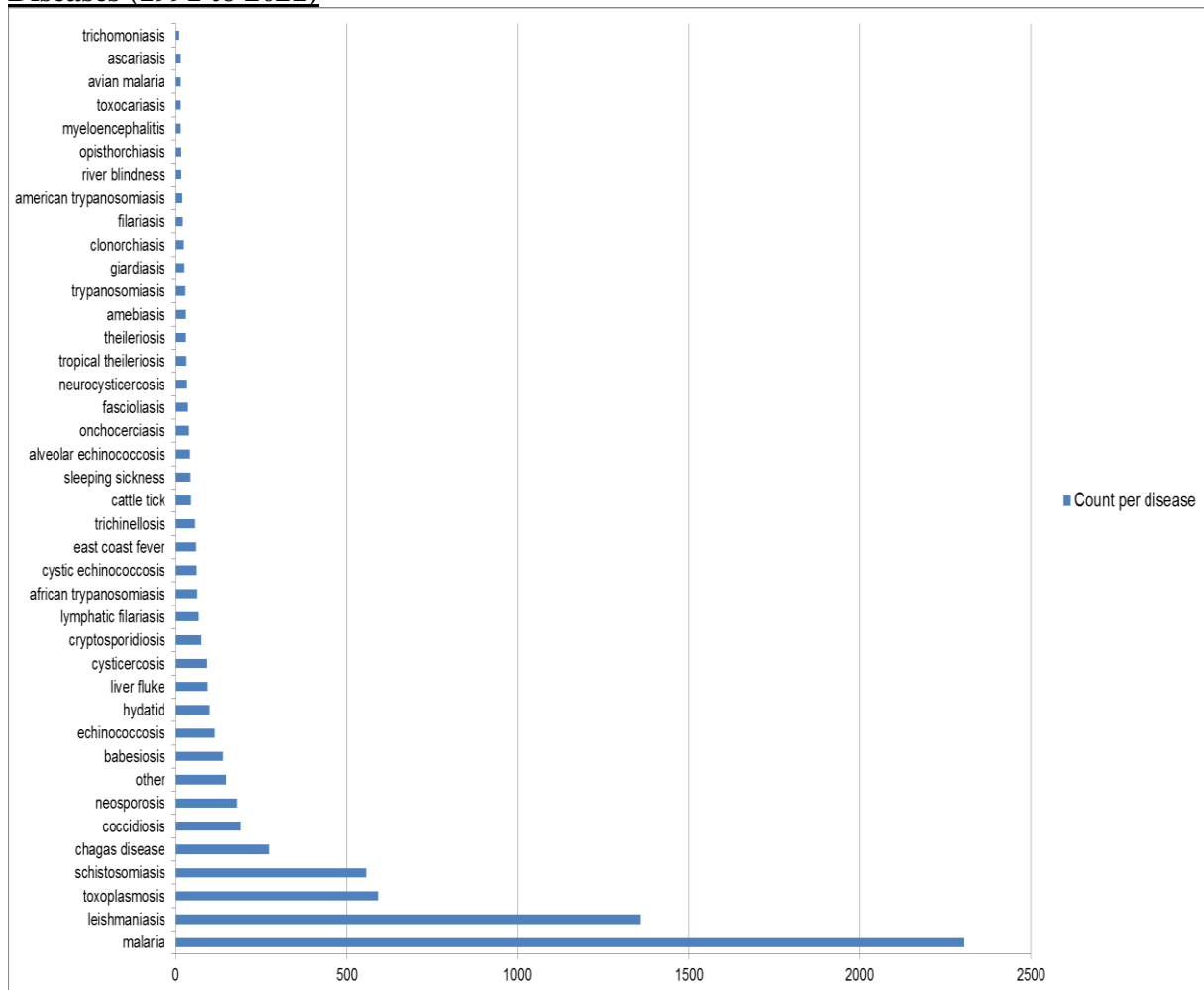
Animal models (1991 to 2021)



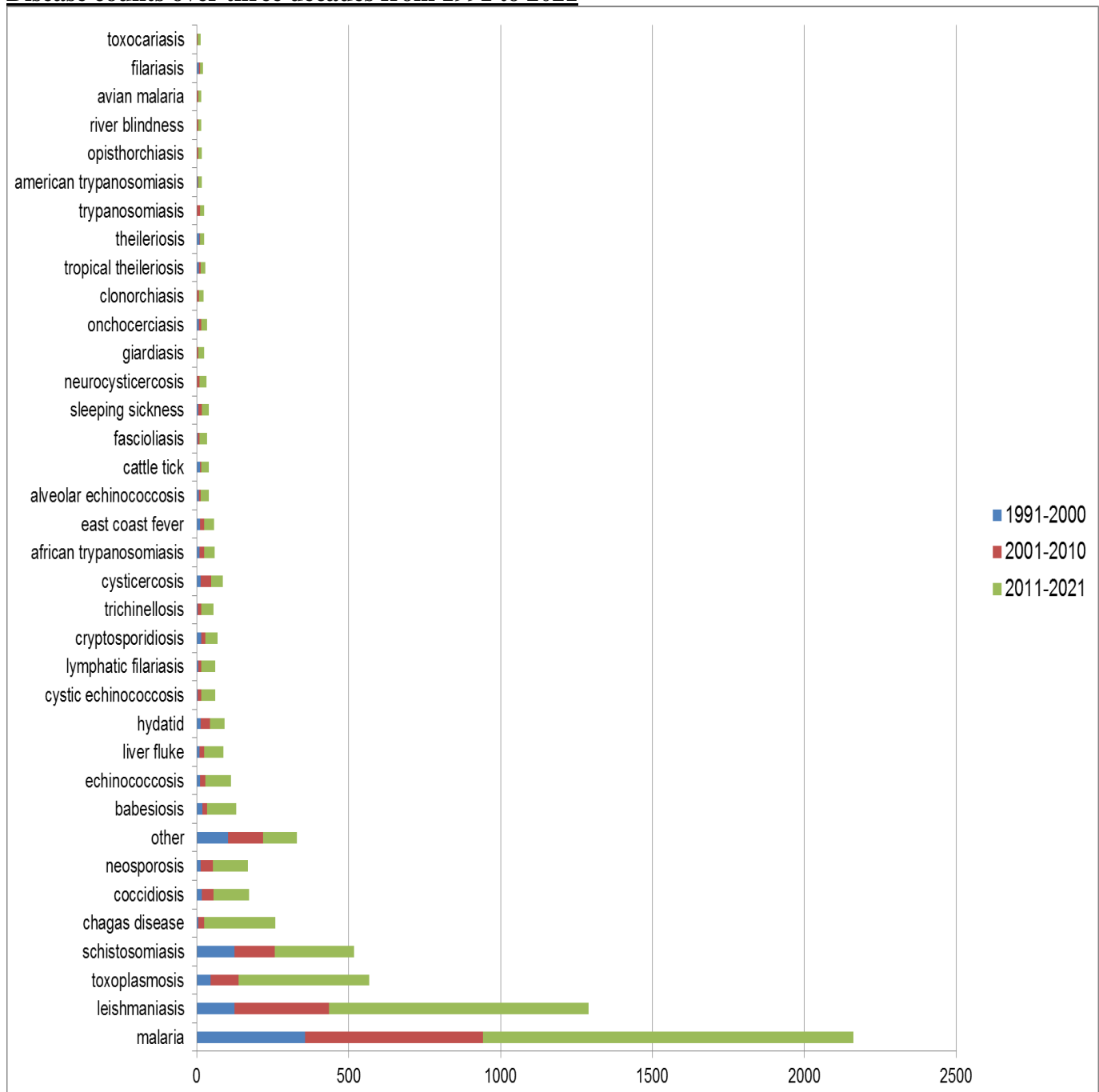
'Animal model' counts over three decades from 1991 to 2021



Diseases (1991 to 2021)



Disease counts over three decades from 1991 to 2021



Genus (1991 to 2021)

